

## MIT Open Access Articles

*Sweep distortion removal from  
terahertz images via blind demodulation*

The MIT Faculty has made this article openly available. **Please share**  
how this access benefits you. Your story matters.

**Citation:** Aghasi, Alireza et al., "Sweep distortion removal from terahertz images via blind demodulation." *Optica* 3, 7 (July 2016): 754-62 doi. 10.1364/OPTICA.3.000754 ©2016 Authors

**As Published:** <https://dx.doi.org/10.1364/OPTICA.3.000754>

**Publisher:** The Optical Society

**Persistent URL:** <https://hdl.handle.net/1721.1/128778>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# Sweep distortion removal from terahertz images via blind demodulation

ALIREZA AGHASI,<sup>1,\*</sup> BARMAK HESHMAT,<sup>1</sup> ALBERT REDO-SANCHEZ,<sup>1</sup> JUSTIN ROMBERG,<sup>2</sup> AND RAMESH RASKAR<sup>1</sup>

<sup>1</sup>Massachusetts Institute of Technology Media Lab, 75 Amherst Street, E14-374, Cambridge, Massachusetts 02139, USA

<sup>2</sup>School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

\*Corresponding author: [aghasi@mit.edu](mailto:aghasi@mit.edu)

Received 30 March 2016; revised 14 June 2016; accepted 17 June 2016 (Doc. ID 262127); published 7 July 2016

Inspection and imaging through scattering layers is a decades-old problem in optics. Unlike x-ray and ultrasonic imaging techniques, time-domain spectroscopy methods can provide detailed chemical and structural information of subsurfaces along with their depth. Although high-resolution time-of-flight measurement in time-domain spectroscopy provides 3D information, unfortunately it also induces an unwanted sensitivity to misalignments of the system and distortion of the layers themselves. Such high sensitivity to alignment and sample surface is a well known problem in time-domain and interferometric imaging, and is a major concern when the alignment error is comparable to the pulse wavelength. Here, we propose and implement an algorithmic framework based on low-rank matrix recovery and alternating minimization to remove such unwanted distortions from time-domain images. The method allows for recovery of the original sample texture in spite of the presence of temporal-spatial distortions. We address a blind-demodulation problem where, based on several observations of the sample texture modulated by undesired sweep distortions, the two classes of signals are separated with minimal damage to the main features. The performance of the method is examined in both synthetic and real data in the case of a terahertz time-domain system, and the successful reconstructions are demonstrated. The proposed general scheme can be implemented to advance inspection and imaging applications in THz and other time-resolved spectral imaging modalities. © 2016 Optical Society of America

**OCIS codes:** (100.3010) Image reconstruction techniques; (100.6950) Tomographic image processing; (110.4155) Multiframe image processing.

<http://dx.doi.org/10.1364/OPTICA.3.000754>

## 1. INTRODUCTION

Due to its fine time resolution and broad spectral coverage, terahertz time-domain spectroscopy (THz-TDS) has become a leading method in THz spectroscopy [1,2], imaging [3], and non-destructive testing [4] of dielectric structures. There is extensive literature on improving THz imaging capability. In general, a large body of literature is focused on improving the hardware, as THz-TDS power levels are usually at sub-microwatt levels [5–7]. This perspective can be extended to improving THz acquisition methodologies such as compressive [8] and wide-field acquisitions [9–12]. On the signal-processing side, the mainstream research has been focused on improvement of signal-to-noise ratio (SNR) [13,14]. Many of these SNR improvement methods have been developed for transmission mode spectroscopy and have appreciable alignment with infrared spectroscopy [15–17]. However, when it comes to imaging, inspection, and content extraction of surfaces and layered structures, THz-TDS can suffer significantly from phase distortions along the sample surface. These sweeping distortions in THz time-domain imaging appear to be the dominant challenge for in-depth imaging and

content extraction in densely layered structures, irregular surfaces, or any 2D geometry. Unfortunately, these heavy distortions are directly induced by the nature of the sample and THz-TDS measurement scheme, and cannot be addressed by minor hardware improvement, conventional denoising, or SNR enhancement techniques.

In this paper, we propose and demonstrate a mathematical framework that enables the demodulation of the recorded signal from sweep distortions induced by slight depth variations or the layered structure inter-reflections. We view the problem as a demodulation of the distortion profiles from the sample texture, based on the reflected wave measured at different instances of time. The problem of interest is modeled as a bilinear inverse problem (BIP), which can be addressed using a low-rank matrix-recovery framework. We propose an alternating minimization scheme, where a prior structure is considered for each factor in the BIP. More specifically, we assume the distortion profiles belong to a low-dimensional subspace (which can be extracted from the raw data) and the layer structure is of a binary nature. The model considered for the layer texture is in fact a shape-based

approximation (see [18–20] for examples), a phase corresponding to the main texture, and another phase representing the anomalies and inclusions. Using the given priors for each factor in the BIP, we alternatively solve the demodulation problem to exclude the undesired sweep distortions from the THz images.

Our mathematical presentation mainly relies on multidimensional calculus. We use bold characters to denote vectors and matrices. Considering a matrix  $\mathbf{A}$  and the index sets  $\Gamma_1$  and  $\Gamma_2$ , we use  $\mathbf{A}_{\Gamma_1,:}$  to denote the matrix obtained by restricting the rows of  $\mathbf{A}$  to  $\Gamma_1$ . Similarly,  $\mathbf{A}_{:, \Gamma_2}$  denotes the restriction of  $\mathbf{A}$  to the columns specified by  $\Gamma_2$ , and  $\mathbf{A}_{\Gamma_1, \Gamma_2}$  is the submatrix with the rows and columns restricted to  $\Gamma_1$  and  $\Gamma_2$ , respectively.

The remainder of this paper is organized as follows. In Section 2, we overview the physics of the problem and discuss the main demodulation problem to be addressed. In Section 3, we present the main algorithmic framework to address the demodulation problem using an alternating minimization scheme. Finally, in Section 4, the sensitivity and performance of the algorithm are assessed with synthetic data. We further experimentally demonstrate the blind demodulation of the data recorded from both single and multilayered structures and report the algorithm outcomes for the experimental data. We conclude this section with some remarks and future directions of research.

## 2. PHYSICS OF THE PROBLEM

Consider an electric field that is linearly polarized in the  $x$  direction, propagating along the  $z$  direction in the free space. The waveform is considered to be a finite-duration THz pulse  $\chi(t)$ , for which the traveling field along every point  $(x, y)$  is

$$\vec{E}_0^+(z, t) = \chi\left(t - \frac{z}{c}\right) \vec{a}_x$$

and  $c$  is the wave speed. Due to the confocal nature of the measurement at each  $(x, y)$  position, the problem of interest can be considered as extracting the contents of a dielectric slab placed perpendicular to the wave-propagation axis, based on analyzing the reflected field  $\vec{E}_0^-(z, t)$  [Figs. 1(a) and 1(b)].

For a homogeneous layer of width  $d$ , with reflection coefficient  $\rho$  and refraction index  $n_\rho$ , the returned signal can be analytically expressed by convolving the pulse  $\chi(\cdot)$  with a train of impulse functions with decaying coefficients. More specifically [21],

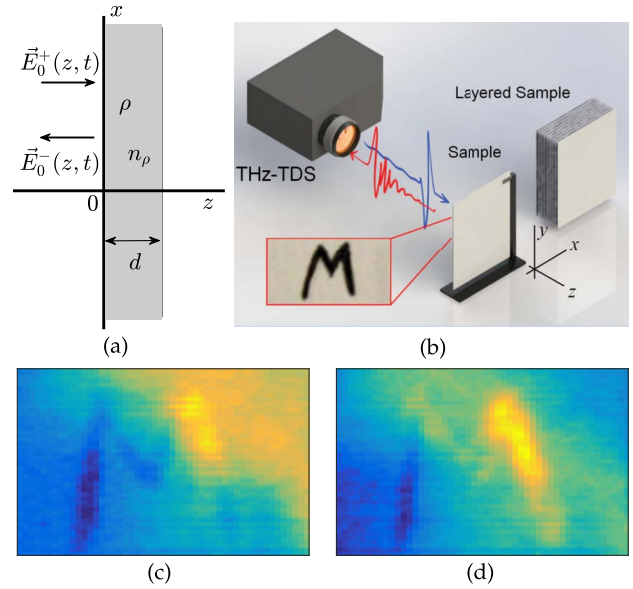
$$\vec{E}_0^-(z, t) = \rho u\left(t + \frac{z}{c}\right) \vec{a}_x, \quad (1)$$

where

$$u(\tau) = \chi(\tau) * \left( \delta\left(\frac{\tau}{\tau_\rho}\right) - \frac{1 - \rho^2}{\rho^2} \sum_{m=1}^{\infty} \rho^{2m} \delta\left(\frac{\tau}{\tau_\rho} - 2m\right) \right), \quad (2)$$

and  $\tau_\rho = n_\rho d / c$ . Each impulse term in Eq. (2) corresponds to a reflection. Particularly, the first term corresponds to the reflection from the front surface of the slab, the second term ( $m = 1$ ) corresponds to the back surface reflection, and the remaining terms correspond to the returned waves after a number of inter-reflections within the slab.

The response in Eq. (1) can still be reasonably accurate when the dielectric slab is homogeneous along the  $z$  axis, i.e.,  $\rho(x, y, z) = \rho_0(x, y)$  for  $0 \leq z \leq d$ . It only requires plugging the point-wise values of  $\rho$  and  $\tau_\rho$  in the formulation.



**Fig. 1.** (a) Dielectric slab, emitted and reflected electrical fields. (b) Setup schematics; blue is the emitted THz field and the red waveform is indicative of a typical reflected signal in one pixel. (c, d) Induction of sweep distortion: observation of the returned field at two different time instances.

In the case of perfect or close to perfect reflection, where  $\rho \approx 1$ , the first impulse term in Eq. (1) dominates all the other terms and the returned wave is a simple modulation of  $\rho$  with the pulse waveform. When  $0 < \rho \ll 1$ , the most dominant terms in  $u(\tau)$  correspond to the first and second impulses and the remaining terms decay exponentially fast. Specifically, in many practical scenarios where the pulse width is sufficiently small and the slab width  $d$  is large, the first reflected waveforms do not overlap with the subsequent reflections. In this case, still a modulation of  $\rho$  with the pulse waveform is observed in different time intervals.

For multilayer dielectric slabs, the impulse response can still be cast as an impulse train; however, deriving closed-form expressions for the coefficients is a sophisticated task and beyond the scope of this paper. When the reflection coefficients of the layers are small (and hence transmission is the leading phenomenon), the dominant terms associated with different layers remain reasonably distinct from one another. If the pulse width is sufficiently small, for each layer the corresponding dominant reflection is still proportional to the reflection coefficient of that layer.

In all the situations just stated, for a measurement location  $(x, y, z_0)$  with a fixed  $z$  component and sampling time  $t = t_0$ , the reflected signal is  $\rho(x, y)u(t_0 + z_0/c)$ , which is a constant multiple of  $\rho(x, y)$ . In other words, the reflected images should contain the same pattern as  $\rho$ . However, due to nonideal configurations such as a tilted or uneven sample, the effective measurement points are  $(x, y, z_0 + \varepsilon(x, y))$ , where  $\varepsilon(x, y)$  is a function of small magnitude—for instance, in the case of a tilted sample  $\varepsilon(x, y) = \alpha_1 x + \alpha_2 y$ , where  $\alpha_1$  and  $\alpha_2$  are small constants.

Technically, the demodulation problem of interest in this paper corresponds to extracting the sample structure  $\rho(x, y)$ , based on observing the reflected signal  $\vec{E}_0^-(x, y, z, t)$  at a fixed  $z$  coordinate and time samples  $t = t_1, \dots, t_M$ . Due to the non-ideal configurations just stated, the returned signal, sampled at a time  $t_j$ , is in the form of  $\rho(x, y)u_j(x, y) + n_j(x, y)$ , where  $u_j(x, y)$

depends on the pulse waveform, sampling time, and  $\varepsilon(x, y)$ . The additive term  $n_j(x, y)$  models the noise uncertainty and undesired electromagnetic interactions. We will refer to the  $u_j(x, y)$  factor as a sweep distortion profile, which is often a slowly varying function when the misconfiguration is small.

Specifically, inspired by the applications presented in this paper, we are interested in single or multilayer composite slabs of a binary nature where, for each layer,

$$\rho(x, y, z) = \begin{cases} \rho^0 & (x, y) \in D_0 \\ \rho^1 & (x, y) \in D_1 \end{cases} \quad (3)$$

and  $\rho^0$  and  $\rho^1$  are constant. An example of a composite slab which roughly follows such a model is a card marked with deep ink [see example in Fig. 1(b)].

Figures 1(c) and 1(d) show examples of the observed reflection from a binary slab (letter “M” printed on a card) at two different time instances. We are willing to make a binary characterization of the sample content by demodulating and excluding the multiplicative sweep distortion profiles affecting the desired image. In the remainder of this section, we will present an inversion scheme to characterize the slab contents based on such multiplicative observations.

### 3. METHODS

#### A. General Setup

For a more concise formulation, we use the more compact spatial notation  $\mathbf{x} = (x, y)$ . Based on the discussions in the previous section, our observation is in the form of

$$y_j(\mathbf{x}) = \rho(\mathbf{x})u_j(\mathbf{x}) + n_j(\mathbf{x}), \quad j = 1, \dots, M, \quad \mathbf{x} \in D, \quad (4)$$

where  $y_j(\mathbf{x})$  represents the  $j$ th observed reflected image dependent on the planar coordinate  $\mathbf{x}$ ,  $\rho(\mathbf{x})$  is the true dielectric texture profile common across all the layers,  $u_j(\mathbf{x})$  is the sweep distortion profile corrupting layer  $j$ ,  $n_j(\mathbf{x})$  models the noise and uncertainty, and  $D$  is the domain of imaging for  $\mathbf{x}$ . As a reasonable model, we consider a normal distribution for the noise, that is,  $n_j(\mathbf{x}) \sim \mathcal{N}(0, \sigma^2)$  are independent and identically distributed (i.i.d) random variables for  $j = 1, \dots, M$  and  $\mathbf{x} \in D$ .

The ultimate goal is recovering the image  $\rho(\mathbf{x})$  and the distortion profile  $u_j(\mathbf{x})$  based on the observation of  $y_j(\mathbf{x})$ . Our strategy to address the problem is considering known subspace models for the signals  $u_j(\mathbf{x})$ . More specifically,  $u_j(\mathbf{x}) \in \mathcal{S}^j$ , where  $\mathcal{S}^j = \text{span}(s_1^j(\mathbf{x}), \dots, s_{N_j}^j(\mathbf{x}))$  and  $N_j$  is the  $j$ th subspace dimension. In Section A.3, we introduce an algorithm to construct a set of low-dimensional subspaces  $\mathcal{S}^1, \dots, \mathcal{S}^M$  from the set of observations  $y_1(\mathbf{x}), \dots, y_M(\mathbf{x})$ . Also, based on the structure of the problem, we may benefit from prior assumptions about  $\rho(\mathbf{x})$ , as will be detailed in the remainder of this section.

#### 1. Problem Reformulation as a Low-Rank Recovery Scheme

Consider discretizing the domain  $D$  into a collection of pixels  $\{\mathbf{x}_i\}_{i=1}^P$ . Based on the model in Eq. (4), a natural way of inverting the data for the image and sweep-distortion profiles is through the following minimization:

$$\min_{\rho, \mathbf{u}_1, \dots, \mathbf{u}_M} \sum_{j=1}^M \|\mathbf{y}_j - \rho \odot \mathbf{u}_j\|_2^2 \quad \text{s.t. } \mathbf{u}_j \in \text{span}(s_1^j, \dots, s_{N_j}^j). \quad (5)$$

Here,  $\odot$  denotes the Hadamard (point-wise) product and the bold scripts are vectors of length  $P$ , corresponding to the variables in Eq. (4). Consider  $\mathcal{S}^j$  to be a matrix representation of the subspace  $\mathcal{S}^j$  and  $\mathbf{Q}$  to be a subspace in which the signal  $\rho$  lives (in the case of no such assumption,  $\mathbf{Q}$  can be simply the identity matrix spanning the canonical basis). Under this assumption, the minimization in Eq. (5) can be cast as

$$\min_{\beta, \alpha_1, \dots, \alpha_M} \sum_{j=1}^M \|\mathbf{y}_j - (\mathbf{Q}\beta) \odot (\mathcal{S}^j \alpha_j)\|_2^2, \quad (6)$$

where the search for the unknown vectors is performed in the corresponding subspaces. Suppose that  $\alpha = [\alpha_1^T, \dots, \alpha_M^T]^T$  and  $\mathbf{P}_j$  is a suitable selection matrix (subset of the rows of the identity matrix) such that  $\alpha_j = \mathbf{P}_j \alpha$ . For the point-wise product of the image and the sweep distortion profiles, we have

$$\begin{aligned} (\rho \odot \mathbf{u}_j)_i &= (\mathbf{Q}\beta)_i \odot (\mathcal{S}^j \alpha_j)_i = (\mathbf{Q}_{i,:} \beta) \mathcal{S}_{i,:}^j \mathbf{P}_j \alpha \\ &= (\mathbf{Q}_{i,:} \beta)^T \mathcal{S}_{i,:}^j \mathbf{P}_j \alpha = \text{tr}(\alpha \beta^T \mathbf{Q}_{i,:}^T \mathcal{S}_{i,:}^j \mathbf{P}_j) \\ &= \langle \beta \alpha^T, \mathbf{Q}_{i,:}^T \mathcal{S}_{i,:}^j \mathbf{P}_j \rangle. \end{aligned} \quad (7)$$

In other words, considering  $\mathbf{X} = \beta \alpha^T$  to be the outer product of the unknown vectors  $\beta$  and  $\alpha$ , the noise-free observations correspond to the inner products of  $\mathbf{X}$  with known matrices  $\mathbf{Q}_{i,:}^T \mathcal{S}_{i,:}^j \mathbf{P}_j$ . Basically, if we define a linear operator

$$(\mathcal{A}(\mathbf{X}))_{ij} = \langle \mathbf{X}, \mathbf{Q}_{i,:}^T \mathcal{S}_{i,:}^j \mathbf{P}_j \rangle, \quad (8)$$

the noise-free observations are linearly dependent on  $\mathbf{X}$  while the original version of the problem in Eq. (6) is nonlinear in terms of  $\beta$  and  $\alpha$ . However, since the outer product of two nonzero vectors is a rank-one matrix, an equivalent reformulation of Eq. (6) is the minimization

$$\min_{\mathbf{X}} \text{rank}(\mathbf{X}) \quad \text{s.t. } \|\mathbf{Y} - \mathcal{A}(\mathbf{X})\|_F \leq \epsilon, \quad (9)$$

where  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M]$  and  $\|\cdot\|_F$  denotes the Frobenius norm.

Despite the linear relationship between the unknown variable and the noise-free observations, the rank constraint in Eq. (9) makes it a combinatorial problem. As a remedy, a well-known approximation strategy is to use the nuclear norm of  $\mathbf{X}$ , denoted by  $\|\mathbf{X}\|_*$ , as a convex proxy to  $\text{rank}(\mathbf{X})$ :

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{s.t. } \|\mathbf{Y} - \mathcal{A}(\mathbf{X})\|_F \leq \epsilon. \quad (10)$$

It is an interesting fact that under certain conditions on  $\mathcal{A}$  (e.g., see [22]), the approximate solution of Eq. (10) can coincide with the solution of Eq. (9). Once a rank-one solution  $\mathbf{X}^*$  is available, the factors  $\alpha$  and  $\beta$  can be determined up to a constant multiple ( $\beta = \mathbf{X}_{:,1}^*$  and  $\alpha^T = \mathbf{X}_{1,:}^* / \mathbf{X}_{1,1}^*$ ).

Other than the relaxation technique, there are other methods reported in the literature to address instances of Eq. (9). We are specifically interested in the alternating minimization approach [23] since it allows the incorporation of prior information into the reconstruction of each bilinear factor.

#### 2. Alternating Minimization: A Maximum a Posteriori Framework

In minimizing the nonconvex program [Eq. (5)], a reliable technique would be to set up an alternating minimization scheme. Basically, by fixing any of the operands in the bilinear term in



Eq. (5) (either  $\rho$  or  $u_j$ ), the problem turns into a standard least squares problem. The process would be to initialize one of the operands in the bilinear model, say  $\rho$ ; with this quantity fixed, the resulting least squares in terms of  $u_j$  is solved. By plugging in the acquired solutions for  $u_j$ 's, we can solve another least squares problem in terms of  $\rho$  and proceed with such alternation until convergence.

This alternation approach is capable of approximating the solution to Eq. (9). Again, under certain incoherence conditions on  $\mathcal{A}$  (slightly stricter than the ones stated for nuclear norm minimization [22,23]), the approximate solution coincides with the solution of Eq. (9). The major advantage of using an alternation scheme over the nuclear norm surrogate is the possibility of incorporating prior information beyond the subspace constraint on each factors of the problem in Eq. (4) (e.g., see [24]).

Basically, the proposed scheme corresponds to an alternation between the maximum likelihood (ML) estimates of  $\rho$  and  $u_j$ . When some level of prior knowledge about the  $\rho$  and/or  $u_j$  exists, the framework could be generalized to an alternation between the maximum *a posteriori* (MAP) estimates of  $\rho$  and  $u_j$ , as will be detailed in the remainder of this section.

Consider  $f(\rho, \{u_j\}_{j=1}^M | \{y_j\}_{j=1}^M)$  to be the joint probability density function (pdf) of  $\rho$  and  $u_j$  given the measurements  $y_j$ . The MAP estimates of  $\rho$  and  $u_j$  correspond to the maximizing parameters of the posterior distribution:

$$\{\rho^{\text{MAP}}, \{u_j^{\text{MAP}}\}_{j=1}^M\} = \underset{\rho, u_1, \dots, u_M}{\operatorname{argmax}} f(\rho, \{u_j\}_{j=1}^M | \{y_j\}_{j=1}^M). \quad (11)$$

As a natural generalization, the alternating MAP estimation corresponds to the process sketched in Algorithm 1.

#### Algorithm 1: MAP Alternation Scheme

```

1:  $k \leftarrow 0$ , initialize  $\rho^{(0)}$ 
2: repeat
3:    $\{u_j^{(k)}\}_{j=1}^M = \operatorname{argmax}_{u_1, \dots, u_M} f(\{u_j\}_{j=1}^M | \rho^{(k)}, \{y_j\}_{j=1}^M)$ 
4:    $\rho^{(k+1)} = \operatorname{argmax}_{\rho} f(\rho | \{u_j^{(k)}\}_{j=1}^M, \{y_j\}_{j=1}^M)$ 
5:    $k \leftarrow k + 1$ 
6: until convergence

```

In the remainder of this section, we elaborate on how to implement the proposed scheme using certain facts about the problem. We would like to note that we make some reasonable assumptions and provide some slight modifications in implementing the steps outlined in Algorithm 1.

#### 3. Iterative Reconstruction of the Sweep Distortion Profiles

In this section, we mainly focus on addressing the first stage of Algorithm 1 (line 3), which is determining the distortion profiles  $u_j$  based on the observations  $\{y_j\}_{j=1}^M$  and the knowledge of  $\rho$ .

To address this problem we only assume that, based on the physics of the problem, the signals  $u_j$  reside in a low-dimensional subspace. In other words,  $u_j = \mathcal{S}^j \alpha_j$  for the discrete representation. However, we make no prior assumptions about the coefficient vector  $\alpha_j$ . Since  $\rho \odot u_j = \operatorname{diag}(\rho) u_j$ , we simply consider solving the least squares problems of the form

$$\alpha_j^{(k)} = \underset{\alpha}{\operatorname{argmin}} \|y_j - \operatorname{diag}(\rho^{(k)}) \mathcal{S}^j \alpha\|_2^2. \quad (12)$$

The solution to Eq. (12) is

$$\alpha_j^{(k)} = (\operatorname{diag}(\rho^{(k)}) \mathcal{S}^j)^\dagger y_j,$$

where  $\mathcal{A}^\dagger$  denotes the pseudoinverse of  $\mathcal{A}$ . In a sense, by making no assumptions about  $\alpha_j$ , instead of performing a MAP estimation (as outlined in Algorithm 1) for a given  $\rho^{(k)}$ , we are performing an ML estimation of the sweep distortion profiles by solving the least squares problems [Eq. (12)].

The choice of subspace plays a key role for this problem. In the remainder of this subsection, we present a one-time process which generates embedding subspaces  $\mathcal{S}^j$  based on the observations  $y_j$ .

From a technical standpoint,  $y_j(\mathbf{x}) \approx \rho(\mathbf{x}) u_j(\mathbf{x})$  where, as stated earlier,  $u_j$  is a rather smooth function and, for our application,  $\rho(\mathbf{x})$  is a function of almost binary nature. Therefore, if we have a multi-scale representation of  $y_j$ , there should be a good overlap between  $u_j$  and the course-level approximation of  $y_j$ . The high-frequency components of  $y_j$  should be mainly in hold of  $\rho$  and the noise.

Based on this argument, consider a wavelet representation of the signal  $y_j$  as

$$y_j(\mathbf{x}) = \sum_w c_{\ell_0, w} \phi_{\ell_0, w}(\mathbf{x}) + \sum_{\ell=\ell_0}^{\infty} \sum_w d_{\ell, w} \psi_{\ell, w}(\mathbf{x}),$$

where  $\phi_{\ell_0, w}$  are the scaling functions,  $\psi_{\ell, w}$  is the wavelet basis, and  $\ell_0$  is a fixed scaling level [25]. For a fixed subspace order  $N_j$ , our proposed subspace selection for  $u_j$  corresponds to the top  $N_j$  scaling/wavelet basis functions with the largest coefficients (in magnitude). In other words, we sort the wavelet coefficients in descending magnitude order and select the basis functions associated with the top  $N_j$  coefficients. The selected basis (in discrete form) corresponds to the columns of  $\mathcal{S}^j$ . This process is performed only once during the entire reconstruction and once the subspaces are determined, they remain fixed throughout the alternation scheme.

#### 4. Iterative Reconstruction of the Image Profile

Inspired by the second stage proposed in Algorithm 1, the main objective in this section is finding the MAP estimate for  $\rho(\mathbf{x})$ . In order to maintain a more compact formulation and closed-form expressions, we make some independence assumptions about the pixel values as will be detailed in the remainder of this section.

Based on the binary model considered, we assume that each pixel in the image  $\rho$  belongs to one of the two classes 0 or 1, and denote the pixel class by  $\mathcal{C}(\mathbf{x})$ . Following the categorization in Eq. (3), the pixel values in class 0 concentrate around  $\rho^0$  and the pixel values associated with class 1 concentrate around  $\rho^1$ . Consider  $\rho = [\rho_1, \dots, \rho_P]^T$  and  $\mathcal{C} = [\mathcal{C}_1, \dots, \mathcal{C}_P]^T$  to be vectors containing the pixel values and the class labels for the discrete image.

For our estimation problem, we are willing to address the maximization problem

$$\max_{\rho, \mathcal{C}} G(\rho, \mathcal{C}),$$

where

$$G(\rho, \mathcal{C}) \triangleq f(\rho, \mathcal{C} | \{u_j^{(k)}\}_{j=1}^M, \{y_j\}_{j=1}^M). \quad (13)$$

In other words, given the observation  $y_j$  and the distortion profiles  $u_j^{(k)}$ , we are willing to acquire MAP estimates of the image profile value at each pixel along with the binary class that pixel belongs to.

Generally speaking, for arbitrary random variables  $\rho$ ,  $y$ ,  $u$ , and  $\mathcal{C}$  (discrete), using the Bayes' rule we have

$$\begin{aligned} f(\rho, \mathcal{C}|y, u) &= \frac{f(y|\rho, \mathcal{C}, u)f(\rho, \mathcal{C}, u)}{f(y|u)f(u)} \\ &= \frac{f(y|\rho, \mathcal{C}, u)f(\rho|\mathcal{C}, u)}{f(y|u)}\mathbb{P}(\mathcal{C}|u), \end{aligned} \quad (14)$$

where we used  $f(\mathcal{C}, u) = \mathbb{P}(\mathcal{C}|u)f(u)$  to derive the second equality. Based on this result, we have

$$G(\rho, \mathcal{C}) \propto f(\{y_j\}_{j=1}^M|\rho, \mathcal{C})f(\rho|\mathcal{C})\mathbb{P}(\mathcal{C}), \quad (15)$$

where, for the sake of convenience, we dropped the prior knowledge of  $\{u_j^{(k)}\}_{j=1}^M$  in all the expressions. Basically, in Eq. (13), the vectors  $\{u_j^{(k)}\}_{j=1}^M$  are known and deterministic and our intention to derive Eq. (14) is to make the MAP estimation in terms of probability density functions which all make such prior assumption. Also, since the denominator term in Eq. (14) is independent of  $\rho$  and  $\mathcal{C}$ , in Eq. (15) we neglected the corresponding term as it will be a constant factor in the maximization.

To further proceed with simplifying the MAP objective, we assume that the pixel values in  $\rho$  are independent of each other. Based on this assumption and the i.i.d nature of the noise, we can state that

$$G(\rho, \mathcal{C}) \propto \prod_{i=1}^P f(\rho_i|\mathcal{C}_i)\mathbb{P}(\mathcal{C}_i) \prod_{j=1}^M f(y_{j,i}|\rho_i, \mathcal{C}_i), \quad (16)$$

where  $y_{j,i}$  is the  $i$ th element of  $y_j$ . The point-wise distributions appearing in Eq. (16) can now be modeled based on the problem setup.

Clearly, based on Eq. (4),  $f(y_{j,i}|\rho_i, \mathcal{C}_i) \sim \mathcal{N}(\rho_i u_{j,i}^{(k)}, \sigma^2)$ , which is an immediate result of the normal noise model. We also assume that

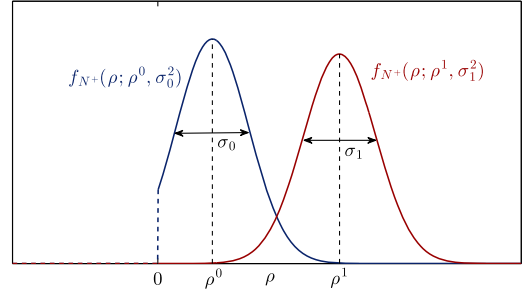
$$\mathbb{P}(\mathcal{C}_i) = p_{\mathcal{C}_i} = \begin{cases} p_0 & \text{if } \mathcal{C}_i = 0 \\ p_1 & \text{if } \mathcal{C}_i = 1 \end{cases}$$

where  $p_0$  and  $p_1 = 1 - p_0$  are somehow the prior estimates of the portion of pixels belonging to each class (e.g., knowing that roughly 20% of the image pixels correspond to the anomaly). If no such information is available *a priori*, a reasonable assumption is  $p_0 = p_1 = 0.5$ .

Finally, with reference to the term  $f(\rho_i|\mathcal{C}_i)$ , knowing that a pixel belongs to class 0 (or 1), we know that its value is likely to be close to  $\rho^0$  (or  $\rho^1$ ). In practice,  $f(\rho_i|\mathcal{C}_i)$  should model the uncertainty on how the values of each class concentrate around the mean class value. While we can use different distributions to model such concentration, for simplicity and in order to obtain closed-form expressions, we assume that  $f(\rho_i|\mathcal{C}_i = 0)$  and  $f(\rho_i|\mathcal{C}_i = 1)$  are in the form of *truncated normal distributions*, taking positive arguments and mainly concentrating around  $\rho^0$  and  $\rho^1$ , respectively. By definition, a random variable  $\rho$  truncated to values greater than zero takes the pdf

$$f_{N^+}(\rho; \tilde{\mu}, \tilde{\sigma}) = \begin{cases} \frac{\gamma}{\tilde{\sigma}} \exp\left(-\frac{(\rho-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right) & \rho \geq 0 \\ 0 & \rho < 0 \end{cases}. \quad (17)$$

Here,  $\tilde{\mu}$  and  $\tilde{\sigma}^2$  are roughly the mean and variance and  $\gamma$  is a normalizing factor to assure the pdf integrates to 1. To model an almost binary nature of the pixel values, we



**Fig. 2.** Distributions for  $f(\rho|\mathcal{C} = 0)$  and  $f(\rho|\mathcal{C} = 1)$  as truncated normal distributions.

assume  $f(\rho_i|\mathcal{C}_i = 0) = f_{N^+}(\rho_i; \rho^0, \sigma_0^2)$  and  $f(\rho_i|\mathcal{C}_i = 1) = f_{N^+}(\rho_i; \rho^1, \sigma_1^2)$  or, more concisely,

$$f(\rho_i|\mathcal{C}_i) = f_{N^+}(\rho_i; \rho^{\mathcal{C}_i}, \sigma_{\mathcal{C}_i}^2).$$

Figure 2 shows the underlying truncated distributions. The positivity assumption in Eq. (17) is imposed by the physics of the problem, where the reflection coefficient of the slab is always considered to be a positive quantity.

Having the constituting terms modeled in Eq. (16), we can proceed with the pixel value and class label MAP estimation. To maximize  $G(\rho, \mathcal{C})$ , we may minimize the negative log function

$$g(\rho, \mathcal{C}) = -\log G(\rho, \mathcal{C}).$$

Based on the proposed distribution models, one can easily verify that

$$g(\rho, \mathcal{C}) = \begin{cases} K + \sum_{i=1}^P g(\rho_i, \mathcal{C}_i) & \rho_i \geq 0 \\ +\infty & \rho_i < 0 \end{cases},$$

where

$$g(\rho_i, \mathcal{C}_i) \triangleq \log \frac{\sigma_{\mathcal{C}_i}}{p_{\mathcal{C}_i}} + \frac{(\rho_i - \rho^{\mathcal{C}_i})^2}{2\sigma_{\mathcal{C}_i}^2} + \sum_{j=1}^M \frac{(y_{j,i} - \rho_i u_{j,i}^{(k)})^2}{2\sigma^2}, \quad (18)$$

and  $K$  is a constant in terms of  $\sigma$ ,  $\gamma_i$ , and the constant factor in Eq. (15). The minimization objective  $g(\rho, \mathcal{C})$  is separable in terms of the cost for each pixel and we can minimize each term individually. Also, minimizing  $g(\rho_i, \mathcal{C}_i)$  can be performed in a serial manner by first minimizing with respect to  $\rho_i$  and then with respect to the class label  $\mathcal{C}_i$ . More concretely,

$$\begin{aligned} \min_{\mathcal{C}_i, \rho_i \geq 0} g(\rho, \mathcal{C}) &= K + \sum_{i=1}^P \min_{\mathcal{C}_i, \rho_i \geq 0} g(\rho_i, \mathcal{C}_i) \\ &= K + \sum_{i=1}^P \min_{\mathcal{C}_i} \min_{\rho_i \geq 0} g(\rho_i, \mathcal{C}_i). \end{aligned} \quad (19)$$

For a fixed given label  $\mathcal{C}_i = \mathcal{C}$ , we can minimize  $g(\rho_i, \mathcal{C})$  with respect to  $\rho_i$  by setting  $\partial g(\rho_i, \mathcal{C})/\partial \rho_i$  to zero. Based on the formulation in Eq. (18), a simple calculation yields

$$\arg\min_{\rho_i} g(\rho_i, \mathcal{C}) = \frac{\sigma^2 \rho^{\mathcal{C}} + \sigma_{\mathcal{C}}^2 \sum_{j=1}^M y_{j,i} u_{j,i}^{(k)}}{\sigma^2 + \sigma_{\mathcal{C}}^2 \sum_{j=1}^M u_{j,i}^{(k)2}}. \quad (20)$$

The function  $g(\rho_i, \mathcal{C})$  is convex in terms of  $\rho_i$ , and the solution to the constrained problem ( $\rho_i \geq 0$ ) is the expression in Eq. (20) if feasible, or 0 otherwise. Basically,

$$\rho_i^* \triangleq \underset{\rho_i \geq 0}{\operatorname{argmin}} g(\rho_i, C_i) = \max \left( \frac{\sigma^2 \rho^C + \sigma_C^2 \sum_{j=1}^M y_{j,i} u_{j,i}^{(k)}}{\sigma^2 + \sigma_C^2 \sum_{j=1}^M u_{j,i}^{(k)2}}, 0 \right). \quad (21)$$

We are now only left with minimizing  $g(\rho_i^*, C_i)$  with respect to  $C_i$ . Since the class labels  $C_i$  only take binary values, we can conveniently find the corresponding minimizer through the following binary comparison:

$$C_i^* \triangleq \underset{C_i \in \{0,1\}}{\operatorname{argmin}} g(\rho_i^*, C_i) = \begin{cases} 1 & \text{if } g(\rho_i^*, 1) < g(\rho_i^*, 0) \\ 0 & \text{if } g(\rho_i^*, 1) \geq g(\rho_i^*, 0) \end{cases}.$$

Equation (20) in some sense generates the optimal solution  $\rho_i$ , but leaves us with an ambiguity about the class label. The label can be obtained based on the comparison just shown to find the ultimate MAP estimator for the image profile. Algorithm 2 summarizes the steps sketched in this section and presents the overall process to perform the demodulation task.

#### Algorithm 2: Decoupling algorithm

---

```

1: input  $\{y_j\}_{j=1}^M, \{N_j\}_{j=1}^M, \rho^0, \rho^1, \sigma_0^2, \sigma_1^2, \sigma^2$ 
2: calculate  $\{S^j\}_{j=1}^M$  from  $\{y_j\}_{j=1}^M$   $\triangleright$  As noted in Section A.3
3:  $\rho \leftarrow 1$ 
4: repeat
5:   for  $j = 1, \dots, M$  do
6:      $\alpha_j \leftarrow (\operatorname{diag}(\rho) S^j)^\dagger y_j$ 
7:      $u_j \leftarrow S^j \alpha_j$ 
8:   for  $i = 1, \dots, P$  do
9:      $w_0 \leftarrow \max \left( \frac{\sigma^2 \rho^0 + \sigma_0^2 \sum_{j=1}^M y_{j,i} u_{j,i}}{\sigma^2 + \sigma_0^2 \sum_{j=1}^M u_{j,i}^2}, 0 \right)$ 
10:     $w_1 \leftarrow \max \left( \frac{\sigma^2 \rho^1 + \sigma_1^2 \sum_{j=1}^M y_{j,i} u_{j,i}}{\sigma^2 + \sigma_1^2 \sum_{j=1}^M u_{j,i}^2}, 0 \right)$ 
11:    if  $g(w_0, 0) \leq g(w_1, 1)$  then  $\triangleright g(\cdot, \cdot)$  is defined in (18)
12:       $\rho_i \leftarrow w_0$ 
13:    else
14:       $\rho_i \leftarrow w_1$ 
15:  until convergence
16: return  $\rho, \{u_j\}_{j=1}^M$ 

```

---

While the parameters  $\sigma_0$  and  $\sigma_1$  have statistical interpretations, from an optimization standpoint they can be considered as free parameters controlling the algorithm's performance. When  $\sigma_0$  and  $\sigma_1$  are set to be small quantities, the algorithm converges faster at the expense of more sensitivity to the initialization (possibility of recovering a local minimizer). Assigning relatively larger values to these quantities paves the path toward identifying the global minimizer in a greater number of iterations. This is also a more reliable parameter selection in the case of noisy observations.

## 4. SIMULATION AND EXPERIMENTS

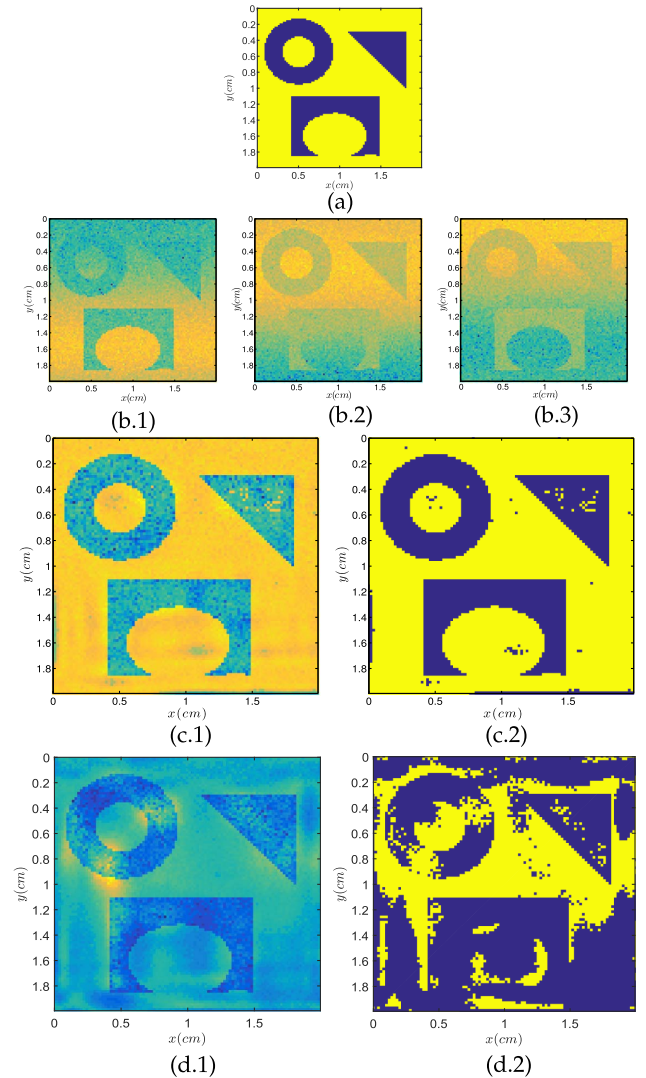
In this section, we assess the performance of the proposed technique in demodulating simulated and real data. The simulation results mainly highlight the performance of the method for various numbers of frames and noise levels in the data. The second set of experiments demonstrates the method's success in removing multiplicative sweep distortion profiles from actual THz measurements.

### A. Simulated Data

We simulate the reflected signal from a dielectric slab with the  $x$ - $y$  profile depicted in Fig. 3(a). The dielectric width is  $d = 100 \mu\text{m}$ . For the binary dielectric slab, the reflection coefficients are  $\rho^0 = 0.3$  and  $\rho^1 = 0.1$ . The sample is probed with a bipolar THz pulse (simply derivative of a Gaussian) as

$$\chi(t) = (t_0 - t) \exp \left( -\frac{(t - t_0)^2}{2T^2} \right) \quad t \geq 0,$$

where  $t_0 = 1$  ps and  $T = t_0/4$ . To model a sweep-modulated signal, we assume that instead of point-wise observations at a constant  $z = z_0$ , the measurements are performed at  $z = z_0 + \alpha_1 x + \alpha_2 y$ , where  $(x, y)$  is the planar coordinate and  $\alpha_1 = 10^{-6}$ ,  $\alpha_2 = 10^{-4}$  are small constants. We take  $M = 10$  uniform samples of the reflected wave in a time period of 0.8 ps. Among the recorded frames in Fig. 3(b) are three sample images where an almost vertically moving sweep distortion across the frames is



**Fig. 3.** Demodulation of simulated data. (a) Reference profile of  $\rho(\mathbf{x})$ . (b.1)–(b.3) Observed reflections modulated by horizontal sweep profiles. (c.1) Demodulated image using the proposed algorithm using  $\sigma_0 = \sigma_1 = 10^{-5}$ . (c.2) Binary approximation of the result in (c.1). (d.1) Recovered image using the nuclear norm minimization. (d.2) Binary approximation of the result in (d.1).



observable. The samples are synthetically corrupted with white noise (SNR = 10 dB).

In recovering the binary profile, for the subspace construction step associated with this example we use  $N_1 = \dots = N_{10} = 100$  most dominant wavelet basis. We use the symlet wavelet family consistently through this example and the remaining experiments.

Figure 3(c.1) reports the reconstructed reflection profile  $\rho$  using our proposed scheme. The algorithm converges in less than 20 iterations and, as may be observed, the result is no more in hold of the distortion traces. A simple rounding of the result to the closest binary value (between  $\rho^0$  and  $\rho^1$ ) yields Fig. 3(c.2), which is reasonably close to the original profile depicted in Fig. 3(a).

Figure 3(d.1) reports the reconstruction using the nuclear norm framework proposed in Section A.1. While some level of distinction between the two binary phases is achieved, the fact that our proposed scheme efficiently uses the binary prior helps us outperform the nuclear norm reconstruction. A similar binary approximation of the nuclear norm reconstruction is demonstrated in Fig. 3(d.2), which clearly fails to characterize some of the main details in the original binary profile.

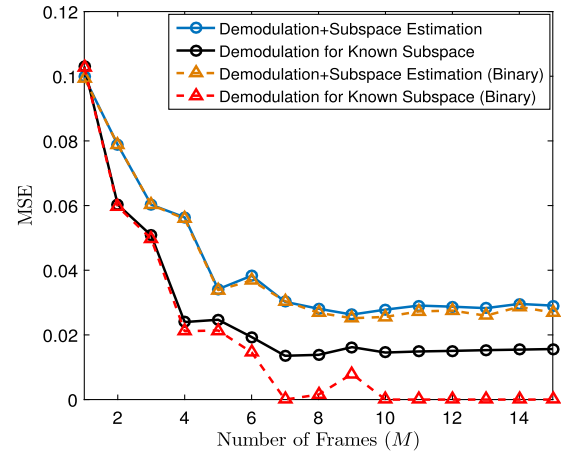
To more extensively analyze the performance of our algorithm, we proceed by the sensitivity analysis of the main parameters affecting the reconstruction. For this purpose, two main scenarios are considered. First, an ideal case where the exact sweep distortion subspace is known *a priori*. From a theoretical standpoint, one way to characterize a subspace for the sweep distortion profiles is to run exactly the same experiment with a homogeneous slab and take the observations as the subspace basis. While experimentally such accurate characterization of the subspace is hard, here we present it as a hypothetical case for comparison purposes and to compare the results against an ideal setup.

For the second class of experiments, we use the proposed wavelet thresholding framework to determine the distortion subspaces. For each proposed scenario, we test the algorithm for various values of  $M$  (number of available frames) and the SNR in the observations. Both the reconstructed  $\rho$  as well as a binary rounded version are compared with the reference image depicted in Fig. 3(a), and the mean squared error (MSE) is reported.

Figure 4 shows the MSE values for various numbers of available frames and a fixed SNR value of 10 dB. The dark solid curve corresponds to the case of exact subspace characterization and the dashed curve standing close to it is the MSE value for the reconstructed  $\rho$  rounded to the closest binary value. We can see that for only 7 frames (or more), a perfect or close to perfect recovery of the binary profile is possible. For fewer numbers of available frames, the MSE still remains at a controlled level.

The lighter solid line (and the dashed line standing close to it) report the MSE values of similar experiments when the sweep distortion subspaces are characterized through the wavelet thresholding. We can see that, aside from a slight performance gap, the MSE values follow a similar reduction pattern as the ideal case. Based on the reported MSE values, the recovery is still close to the reference image. The slight performance gap between this case and the ideal case is the result of error in exact characterization of the sweep distortion subspaces.

Figure 5 shows the MSE values for the case of a fixed number of frames ( $M = 10$ ) and varying observation noise. For this experiment, in the case of known distortion subspace, a perfect recovery of the sweep distortion profile is possible for SNR values

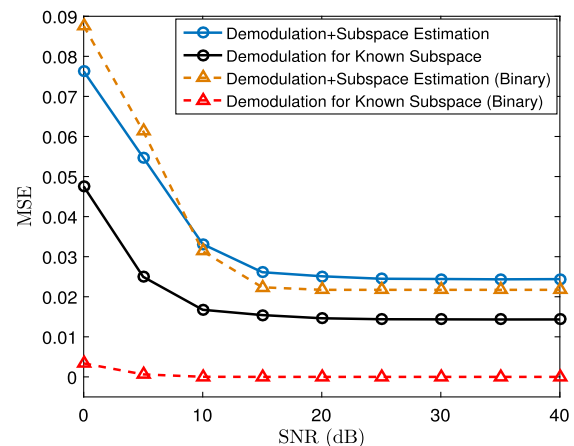


**Fig. 4.** MSE as a function of available number of frames,  $M$ , for various scenarios: a total of 20 frames using uniform sampling in time is made available through the simulation. For each MSE report,  $M$  frames are randomly selected from the 20 frames and passed to the algorithm. The process is performed multiple times for each  $M$  and the average MSE is reported.

approximately exceeding 10 dB (by rounding to the closest binary value). In the case of characterizing the sweep distortion subspace through wavelet thresholding, the MSE values follow a similar trend as the case of knowing the distortion subspace *a priori*. The only difference is a slight performance gap which is again linked to the error in exact characterization of the sweep distortion subspaces.

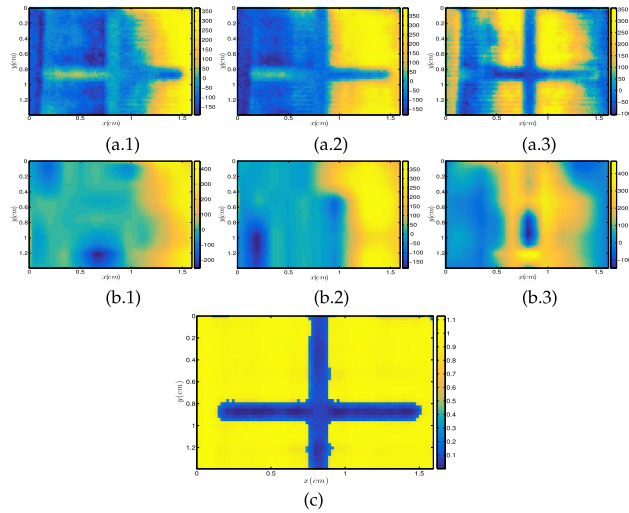
## B. Experimental Data

To experimentally apply the demodulation process, we used a standard THz-TDS system in reflection geometry. We exploited a bipolar pulse with 2.0 THz bandwidth and an effective center frequency of 1.0 THz. In the first experiment (Fig. 6), we used a binary sample with metallic surface and a cross-shaped pattern carved into a 15 mm × 15 mm metal surface with 2 mm depth. As noted in Figs. 6(a.1)–6(a.3), the spatial sweep distortions are dominantly present in the time domain observations. For this experiment,  $M = 25$  frames were available, three of which are shown here.



**Fig. 5.** MSE as a function of SNR for a fixed number of available frames ( $M = 10$ ).





**Fig. 6.** Experimental demonstration of sweep decoupling for a metallic surface. (a.1)–(a.3) Three time instances of the recorded image from the sample; vertical sweep distortions are dominant over the cross pattern. (b.1)–(b.3) Recovered distortion profiles corresponding to the observations shown. (c) Recovered binary profile.

Figures 6(b.1)–6(b.3) and 6(c) report the decoupled sweep distortions from the binary cross profile. We can observe a good characterization of the cross profile thanks to the large number of available frames and high SNR observations. A more challenging experiment with fewer available frames and noisier data is presented next.

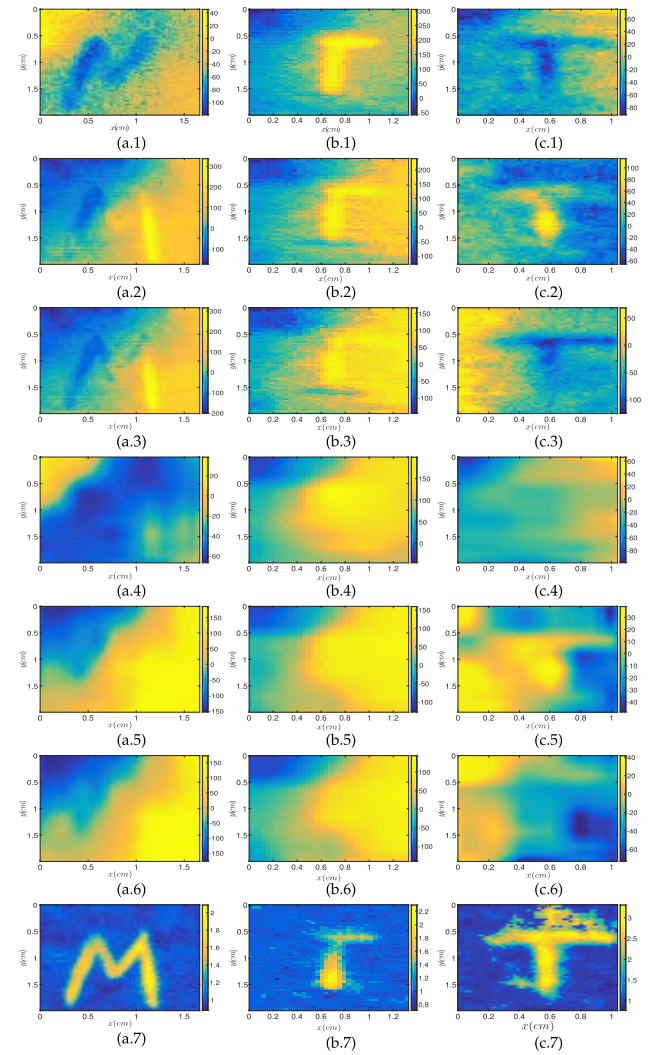
A more challenging experiment uses a three-layer paper sample with the letters M-I-T printed on each page from front to back, respectively. The page dimensions are 15 mm  $\times$  30 mm and the thicknesses are 3.0  $\mu$ m, stacked flush next to each other ( $\sim$ 30  $\mu$ m gap), similar to the structure of a book. Figures 7(a.1)–7(a.3) show the reflected signal from the first layer at three different instances of time. Figures 7(b.1)–7(b.3) and 7(c.1)–7(c.3) show reflection samples from the second and third layers, respectively. The available number of frames corresponding to the first, second, and third layer are  $M = 25, 6, 8$ , respectively.

Followed by the observations in each column of Fig. 7, the demodulation results are presented for the three letters at different pages of the sample, as the THz pulse travels deeper into the pages. The rows 4–6 in Fig. 7 correspond to the sweep distortion profiles decoupled from the observed data. The last row corresponds to the recovered character profiles.

For this multilayer setup, the SNR drops after the first layer and a larger number of inter-reflections is induced; thus, the recovered letters “I” and “T” in panels Figs. 7(b.7) and 7(c.7) have more noise compared to the recovered letter “M” on the first page [panel Fig. 7(a.7)].

The clean results from the first experiment (Fig. 6) on the metallic surface are anticipated due to high SNR and the perfectly flat surface of the polished steel; however, for the paper, the reflection is rather small ( $<0.15$ ), the distortion contrast is at the same level of the signal contrast itself, and the pages have slight depth variations across them. The results from the paper sample in Fig. 7 indicate the true robustness of the decoupling technique for practical samples.

It is worth noting that the major burden in characterizing the sample inhomogeneity profiles (characters in this example) is the



**Fig. 7.** Experimental demonstration of sweep decoupling for a paper stack sample. (a.1)–(a.3) Three time instances of the recorded image from the first page. (b.1)–(b.3) Second page. (c.1)–(c.3) Third page; irregular distortions are dominantly present. (a.4)–(a.6) The algorithm has demodulated the distortion from the first page; (b.4)–(b.6) for the second page; (c.4)–(c.6) third page. (a.7) Recovered letter “M” on the first page. (b.7) Recovered letter “I” on the second page. (c.7) Recovered letter “T” on the third page.

varying signal level in the observed images. For instance, while the contrast between the character “M” and the background can be visually detected in Fig. 7(a.3), the pixel values drastically vary from one portion of the letter to the other. As a result, characterizing the letter based on the pixel values becomes an inconceivable task. However, once a binary profile is reconstructed through the proposed algorithm, other post-processing schemes can be employed to characterize the reconstructions, especially in the case of the more noisy reconstructions such as Fig. 7(c.7). The interested reader is referred to [26–29], where advanced shape-composition techniques are employed to accurately identify binary inclusions in noisy images. This class of techniques can accurately characterize the inclusion in an image like Fig. 7(c.7). Even when the noise is higher, parts of the object are missing due to signal occlusions, and we have clutter or overlapping characters caused by the inter-reflections from neighboring layers.

## 5. CONCLUSION

The main outcome of this research is the development of a general demodulation scheme to process reflected signals conveniently applicable to THz imaging. While a careful processing of the THz data can produce high-resolution images thanks to its high frequency, dealing with phenomena such as inter-reflections, noise, and modulated distortions is inevitable. We showed that reformulation of the problem as a bilinear inverse problem and making practical assumptions, such as the binary prior, can assist us with a successful inversion of THz measurements in THz-TDS systems.

Using basic statistical assumptions about the image and the observation, we were able to develop an iterative inversion scheme which is computationally efficient, distributable, and scalable to big data sets. In fact, the main computational load at every step of the algorithm is a least squares solve. The remaining computations are in the form of thresholding-type operations.

Since the main objective in many applications (such as water profilometry, extracting coded structures, and extracting subsurface cracks) is a binary characterization of the inhomogeneity, we used a binary prior in the modeling in our samples. The algorithm can be naturally generalized to the case of polyadic priors, where more than two levels are considered for the image profile. Basically, the proposed algorithm can be further modified or combined with post- or pre-processing tools to handle a larger class of imaging applications.

**Funding.** Massachusetts Institute of Technology Media Laboratory Consortia (2746038); National Science Foundation (NSF) (RI 1527181).

## REFERENCES

1. Y.-C. Shen, "Terahertz pulsed spectroscopy and imaging for pharmaceutical applications: a review," *Int. J. Pharm.* **417**, 48–60 (2011).
2. M. Tonouchi, "Cutting-edge terahertz technology," *Nat. Photonics* **1**, 97–105 (2007).
3. C. Seco-Martorell, V. López-Domínguez, G. Arauz-Garofalo, A. Redo-Sanchez, J. Palacios, and J. Tejada, "Goya's artwork imaging with terahertz waves," *Opt. Express* **21**, 17800–17805 (2013).
4. A. Redo-Sanchez, N. Laman, B. Schulkin, and T. Tongue, "Review of terahertz technology readiness assessment and applications," *J. Infrared Millimeter Terahertz Waves* **34**, 500–518 (2013).
5. B. Heshmat, H. Pahlevaninezhad, Y. Pang, M. Masnadi-Shirazi, R. Burton Lewis, T. Tiedje, R. Gordon, and T. E. Darcie, "Nanoplasmonic terahertz photoconductive switch on GaAs," *Nano Lett.* **12**, 6255–6259 (2012).
6. B. Heshmat, H. Pahlevaninezhad, and T. Darcie, "Carbon nanotube-based photoconductive switches for THz detection: an assessment of capabilities and limitations," *IEEE Photon. J.* **4**, 970–985 (2012).
7. B. Heshmat, M. Masnadi-Shirazi, R. B. Lewis, J. Zhang, T. Tiedje, R. Gordon, and T. E. Darcie, "Enhanced terahertz bandwidth and power from GaAsBi-based sources," *Adv. Opt. Mater.* **1**, 714–719 (2013).
8. W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Appl. Phys. Lett.* **93**, 121105 (2008).
9. W. Withayachumnankul and D. Abbott, "Terahertz imaging: compressing onto a single pixel," *Nat. Photonics* **8**, 593–594 (2014).
10. C. Li, J. Grant, J. Wang, and D. R. Cumming, "A nipkow disk integrated with Fresnel lenses for terahertz single pixel imaging," *Opt. Express* **21**, 24452–24459 (2013).
11. H. Shen, L. Gan, N. Newman, Y. Dong, C. Li, Y. Huang, and Y. Shen, "Spinning disk for compressive imaging," *Opt. Lett.* **37**, 46–48 (2012).
12. A. W. Lee and Q. Hu, "Real-time, continuous-wave terahertz imaging by use of a microbolometer focal-plane array," *Opt. Lett.* **30**, 2563–2565 (2005).
13. Y. Chen, S. Huang, and E. Pickwell-MacPherson, "Frequency-wavelet domain deconvolution for terahertz reflection imaging and spectroscopy," *Opt. Express* **18**, 1177–1190 (2010).
14. R. Galvão, S. Hadjiloucas, J. Bowen, and C. Coelho, "Optimal discrimination and classification of THz spectra in the wavelet domain," *Opt. Express* **11**, 1462–1473 (2003).
15. Y. Deng, Q. Sun, F. Liu, C. Wang, and Q. Xing, "Terahertz time-resolved spectroscopy with wavelet-transform," in *3rd International Congress on Image and Signal Processing (CISP)* (IEEE, 2010), Vol. 7, pp. 3462–3464.
16. A. D. Burnett, W. Fan, P. C. Upadhyay, J. E. Cunningham, M. D. Hargreaves, T. Munshi, H. G. Edwards, E. H. Linfield, and A. G. Davies, "Broadband terahertz time-domain spectroscopy of drugs-of-abuse and the use of principal component analysis," *Analyst* **134**, 1658–1668 (2009).
17. X. Yin, B. W.-H. Ng, and D. Abbott, *Terahertz Imaging for Biomedical Applications: Pattern Recognition and Tomographic Reconstruction* (Springer, 2012).
18. A. Aghasi, I. Mendoza-Sanchez, E. L. Miller, C. A. Ramsburg, and L. M. Abriola, "A geometric approach to joint inversion with applications to contaminant source zone characterization," *Inverse Prob.* **29**, 115014 (2013).
19. A. Aghasi, M. Kilmer, and E. L. Miller, "Parametric level set methods for inverse problems," *SIAM J. Imaging Sci.* **4**, 618–650 (2011).
20. E. L. Miller, L. M. Abriola, and A. Aghasi, "Environmental remediation and restoration: hydrological and geophysical processing methods," *IEEE Signal Process. Mag.* **29**(4), 16–26 (2012).
21. R. W. Scharstein, "Transient electromagnetic plane wave reflection from a dielectric slab," *IEEE Trans. Educ.* **35**, 170–175 (1992).
22. B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.* **52**, 471–501 (2010).
23. P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing* (ACM, 2013), pp. 665–674.
24. K. Lee, Y. Wu, and Y. Bresler, "Near optimal compressed sensing of sparse rank-one matrices via sparse power factorization," *arXiv:1312.0525* (2013).
25. G. Strang and T. Nguyen, *Wavelets and Filter Banks* (SIAM, 1996).
26. A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, and R. Raskar, "Terahertz time-gated spectral imaging for content extraction through layered structures," *Nat. Commun.* (to be published).
27. A. Aghasi and J. Romberg, "Convex cardinal shape composition," *SIAM J. Imaging Sci.* **8**, 2887–2950 (2015).
28. A. Aghasi and J. Romberg, "Sparse shape reconstruction," *SIAM J. Imaging Sci.* **6**, 2075–2108 (2013).
29. A. Aghasi and J. Romberg, "Learning shapes by convex composition," *arXiv:1602.07613* (2016).